



**QUEEN'S
UNIVERSITY
BELFAST**

Modelling mortality: Are we heading in the right direction?

O'Hare, C., & Li, Y. (2017). Modelling mortality: Are we heading in the right direction? *Applied Economics*, 49(2), 170-187. <https://doi.org/10.1080/00036846.2016.1192278>

Published in:
Applied Economics

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights

Copyright 2016 Informa UK Limited

This is an Accepted Manuscript of an article published by Taylor & Francis in *Applied Economics* on 7/6/2016, available online:
<http://www.tandfonline.com/10.1080/00036846.2016.1192278>

General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

MODELLING MORTALITY: ARE WE HEADING IN THE RIGHT DIRECTION?

COLIN O'HARE[†] AND YOUWEI LI[‡]

ABSTRACT. Predicting life expectancy has become of utmost importance in society. Pension providers, insurance companies, government bodies and individuals in the developed world have a vested interest in understanding how long people will live for. This desire to better understand life expectancy has resulted in an explosion of stochastic mortality models many of which identify linear trends in mortality rates by time. In making use of such models for forecasting purposes we rely on the assumption that the direction of the linear trend (determined from the data used for fitting purposes) will not change in the future, recent literature has started to question this assumption. In this paper we carry out a comprehensive investigation of these types of models using male and female data from 30 countries and using the theory of structural breaks to identify changes in the extracted trends by time. We find that structural breaks are present in a substantial number of cases, that they are more prevalent in male data than in female data, that the introduction of additional period factors into the model reduces their presence, and that allowing for changes in the trend improves the fit and forecast substantially.

JEL Classification: C51, C52, C53, G22, G23, J11

Keywords and Phrases: Mortality, stochastic models, structural breaks, forecasting.

Date: Latest version: May 16, 2016.

[†]Department of Econometrics and Business Statistics, Monash University, Melbourne, Vic 3800, Australia,

[‡]Corresponding author. School of Management, Queen's University of Belfast, BT9 5EE, Belfast, United Kingdom. Tel.: +44 28 9097 4826; Fax: +44 28 9097 4201. Emails: colin.ohare@monash.edu and y.li@qub.ac.uk.

Acknowledgements: We are grateful to Professor Rob Hyndman for useful comments received on drafts of this paper. Financial support for Li from the Australian Research Council (ARC) under Discovery Grant (DP130103210) is gratefully acknowledged.

1. INTRODUCTION

Over the past recent decades, life expectancy in developed countries has risen to historically unprecedented levels. The prospects of future reductions in mortality rates are of fundamental importance in various areas such as demography, actuarial studies, public health, social insurance planning, and economic policy. Over many years, significant progress has been made in mortality modelling (see, for example, Booth and Tickle, 2008), with the most popular approaches being based on the Lee and Carter (1992) extrapolative model. It describes the time-series movement of age-specific mortality as a function of a latent level of mortality, also known as the overall mortality index, which can be modelled using simple time-series methods. The method was initially used to model mortality in the US, but since then has been applied to many other countries (amongst others see Tuljapurkar and Boe, 1998; Carter and Prskawetz, 2001; Lee and Miller, 2001; Booth *et al.*, 2002; Brouhns *et al.*, 2002; Renshaw and Haberman, 2003 and Koissi *et al.*, 2005). Given the popularity and wide spread use of models such as the Lee Carter model in this paper we consider the appropriateness of it and its several extensions. In particular, we use structural break analysis to identify whether the period effects extracted in these models should indeed be modelling using a random walk with drift time series. The findings are extremely important particularly for those who plan to use such models to make long term forecasts of mortality and longevity.

Some consideration has been given to structural changes in mortality trends in the actuarial literature. Li, Chan, and Cheung (2011) applied a broken trend stationary model to the extracted mortality trend κ_t of the Lee-Carter model using the Zivot and Andrews' (1992) procedure. They applied the model to data from the US and from England and Wales and in each case identified break points in the mid 1970's. Their findings confirmed those of Renshaw and Haberman (2003) who also identified an improvement in fitting if an adjustment in the trend was allowed for at 1975. Coelho and Nunes (2011) repeated this analysis using the tests of Harvey *et al.* (2009) and Harris *et al.* (2009) to identify the presence of and date any structural breaks in the extracted mortality trend κ_t . Their study was wider focusing on 18 different countries in total and focusing on both males and females. Notably they found structural breaks in 16 of the 18 countries for males but in only 5 of the 18 countries for females suggesting that any potential acceleration in mortality improvement has had a greater impact on male mortality

than on female mortality. They also found a range of structural break dates from 1955, for Japanese females, through to the year 2000 for Netherlands males. They also forecast with and without an allowance for the identified structural breaks and in the case of Portugal suggest an increase in life expectancy at birth of just over 2 years (80.9 vs. 78.7) when allowing for the break. van Berkum *et al.* (2014) also consider structural breaks within a selection of time series mortality models. They use Dutch and Belgium male mortality data to test for the presence of and consequent impact of structural breaks on mortality forecasts. It is important to note in each of the cases studied the mortality improvement factor κ_t appears to accelerate after the break suggesting that if there is a structural break identified then the resulting model allowing for this break will predict a higher life expectancy. In particular, using a model which doesn't adequately capture any structural change in the improvement in mortality rates for pricing and reserving may lead to an under provision of reserves or prices.

The study of structural breaks within mortality modelling is still relatively recent considering a limited selection of countries and focusing predominantly on male data.¹ In this paper we contribute to the existing literature by carrying out a comprehensive study of 30 countries' mortality data, considering the Lee-Carter model and a selection of extensions including Cairns, Blake and Dowd (2006), Plat (2009), and O'Hare and Li (2012) and using both male and female data. The purpose of this is three fold: (i) to examine how widespread the presence of structural breaks is within mortality data; (ii) to investigate potential differences between male and female data with regard to structural breaks; and (iii) to examine whether the inclusion of additional age, time or cohort effects has any impact on the presence or not of structural breaks. The proposed methodology is applied to both male and female data over the period 1960-2009 and structural changes in the rate of decline in the overall mortality rate are found in a significant number of countries and in particular for each model considered. Structural breaks also appear to be present in male data more than in female data. Where structural breaks are present then allowing for these changes the mortality improvement trend and improves the overall fit quality.

The paper is organized as follows. Section 2 presents a brief review of extrapolative models such as the Lee-Carter model and its extensions. Section 3 discusses the methodology we use to identify the presence of structural breaks. In section 4 we discuss the data we have used in this study. Section 5 presents the results of our analysis to identify structural breaks and to quantify

¹The interested reader is referred to O'Hare and Li (2014)

the impact we demonstrate the improvement in fitting results with and without allowance for the identified structural changes. We also present in sample forecasting results and out of sample forecasts. Finally, section 6 concludes with some ideas for further research.

2. LEE-CARTER AND ITS VARIANTS

The current leading method for modelling mortality rates is the stochastic extrapolation approach. In this method data is first transformed (by taking natural logarithms) and then analysed using statistical methods to identify and extract patterns. These patterns are then modeled using well known time series approaches. The first and most well known stochastic mortality model of this type is the Lee and Carter (1992) model. Based on US data the model uses a stochastic, time series framework to identify a single period effect pattern in the natural logarithm of mortality rates.² This linear trend over time is extracted and using Box-Jenkins an appropriate ARIMA processes is fitted to the data (a random walk with drift in each case). The random walk with drift is forecast and resulting future mortality rates predicted. Also known as a one factor or one principal component approach the model became a benchmark and underlined a new approach to modelling mortality rates for several reasons: the model has an extremely simple structure and so is very easy to communicate; and the use of the random walk with drift enabled the authors not only to predict the expected future mortality rates but also to visualise the uncertainty associated with the predictions. The **Lee-Carter** model, outlined below includes two age dependent parameters a_x and b_x which respectively represent the intercept and gradient for the log mortality rate at each age and the time or period trend κ_t which is modeled using a random walk with drift:

$$(1) \quad \ln(m_{x,t}) = a_x + b_x \kappa_t + \epsilon_{x,t},$$

where a_x and b_x are age effects and κ_t is a random period effect.

The model is known to be over parameterised and applying the necessary constraints as in the original Lee and Carter (1992) paper the a_x are given by

$$a_x = \frac{1}{N} \sum_{t=1}^N \ln m_{x,t}.$$

²We note that recently there are models focused on differences or ratios in mortality rates in an effort to make the underlying time series components more stationary, see, for example, Haberman and Renshaw (2012, 2013) and Mitchell *et al.* (2013).

In the original paper the bilinear part $b_x \kappa_t$ of the model specification was determined as the first singular component of a singular value decomposition (SVD), with the remaining information from the SVD considered to be part of the error structure. The κ_t were then estimated and refitted to ensure the model mapped onto historic data.

Despite the attractiveness of the models simplicity it has several weaknesses. Among many discussions of the Lee-Carter model, Cairns *et al.* (2006, 2009, and 2011) summarized the main disadvantage of the model as having only one factor, resulting in mortality improvements at all ages being perfectly correlated (trivial correlation structure). They also note that for countries where a cohort effect is observed in the past, the model gives a poor fit to historical data. The uncertainty in future death rates is proportional to the average improvement rate b_x which for high ages can lead to this uncertainty being too low, since historical improvement rates have often been lower at high ages. Also, the model can result in a lack of smoothness in the estimated age effect b_x .

Despite the weaknesses of the Lee-Carter model its simplicity has led to it being taken as a benchmark against which other stochastic mortality models can be assessed. There has been a significant amount of literature developing additions to, or modifications of, the Lee-Carter model. For example Booth *et al.* (2002), Brouhns *et al.* (2002), Lee and Miller (2001), Girosi and King (2005), Delwarde *et al.* (2007) and Renshaw and Haberman (2003, 2006, 2011).

Renshaw and Haberman (2006) modified the Lee-Carter model by simply adding a factor to capture cohort effect. The model does have a much better fit for countries such as the UK where a cohort effect has been identified, however it suffers from a lack of robustness perhaps due to the presence of more than one local maximum in the likelihood function. Among others, for instance Currie (2006) noted that if the model was fitted using data from 1961-2000 then the parameters showed qualitatively different characteristics to those obtained when fitting to data from 1981-2000. Furthermore, as noted by Currie (2006), although the model incorporates the cohort effect, for most of the simulated mortality rates the correlation structure is still trivial with the simulated cohort parameters only being relevant for the higher ages at the far end of the projection. Following this analysis Currie (2006) applied a simplified age-period-cohort model of Clayton and Schifflers (1987) to mortality which removed the robustness problem but at the expense of the fitting quality. Renshaw and Haberman (2011) introduced a further simplification of their model which has been shown to have a better fit to many countries.

Cairns, Blake and Dowd (2006) observed that for England & Wales and United States data, the fitted cohort effect appeared to have a trend in the year of birth. This suggested that the cohort effect was compensating for the lack of a second age-period effect, as well as trying to capture the cohort effect in the data. This led them to introduce a two factor model (**CBD** model hereafter),

$$(2) \quad \text{logit}(q_{x,t}) = \kappa_t^1 + \kappa_t^2(x - \bar{x}) + \epsilon_{x,t},$$

where \bar{x} is the mean age in the sample range and (κ_t^1, κ_t^2) are assumed to be a bivariate random walk with drift. The model fits a logistically transformed initial rate of mortality, $q_{x,t}$, using two factors which are both period factors. There is no cohort effect allowed for however, this was rectified in Cairns *et al.* (2009), namely capturing the cohort effect as an additional effect on top of the two age-period effects. The initial rate of mortality can be related to the central rate of mortality $m_{x,t}$ through: $q_{x,t} = 1 - \exp(-m_{x,t})$. All these models have multiple factors resulting in a non-trivial correlation structure which mirrors the reality that improvements in mortality rates are different for different age ranges. A further adaptation was also created allowing for the cohort effect to diminish over time. The main problem with these models arises from the fact that they were designed for higher ages and so ignored the modelling of mortality at the lower ages (for example the accident hump). Cairns *et al.* (2009) argue that the significant cost associated with mortality is at the older ages and thus their modelling focused on those ages. When using these models for full age ranges, the fit quality is relatively poor and the projections are biologically unreasonable.

Plat (2009) wanted to develop a model which maintained the good aspects of the existing models whilst leaving out the weaker features. The result was a four factor model which took its beginnings from the Lee-Carter model and which added factors to capture the second age-period effect, as per the CBD model and the cohort effect, as per the Renshaw and Haberman (2006) model. The innovation in the Plat model was to then add a further period factor affecting only the lower ages and designed to allow the model to fit to the whole age range. The **Plat** model specification is given by:

$$(3) \quad \ln(m_{x,t}) = a_x + \kappa_t^1 + \kappa_t^2(\bar{x} - x) + \kappa_t^3(\bar{x} - x)^+ + \gamma_{t-x} + \epsilon_{x,t},$$

where the a_x is similar to that of the Lee-Carter model and makes sure that the overall shape of the mortality curve by age is reasonable, the κ_t^1 and κ_t^2 model the mortality rates as in the CBD model and the κ_t^3 models the effects specific to the lower ages only where $(\bar{x} - x)^+$ takes the value $(\bar{x} - x)$ when this is positive and zero otherwise. Finally the γ_{t-x} models the cohort effect.

O'Hare and Li (2012) modified the Plat (2009) model to provide a better fit for a wider age range including ages 5-20. They improved the Plat specification by including a quadratic lower age effect. Their model specification is given below (**OL** model hereafter):

$$(4) \quad \ln(m_{x,t}) = a_x + \kappa_t^1 + \kappa_t^2(\bar{x} - x) + \kappa_t^3((\bar{x} - x)^+ + [(\bar{x} - x)^+]^2) + \gamma_{t-x} + \epsilon_{x,t},$$

where a_x makes sure that the basic shape of the mortality curve over ages is in line with historical observations as in the Lee-Carter model (1) and the κ_t^1 factor represents changes in the level of mortality for all ages. Following the reasoning in Cairns *et al.* (2006), the (long-term) stochastic process for this factor should not be mean reverting. The κ_t^2 factor allows changes in mortality to vary between ages reflecting the historical observation that improvement rates can differ for different age classes and κ_t^3 models the effects specific to the lower age only as in the Plat model (3). The adjusted coefficient of κ_t^3 is designed to capture some of the non-linear effects observed at the lower ages, the “quadratic lower age effect”.

Each of these models relies heavily on the assumption that the extracted period effects, κ_t , in the case of the Lee-Carter model, κ_t^1 and κ_t^2 , in the case of the CBD model, κ_t^1 , κ_t^2 , and κ_t^3 in the case of the Plat and OL models can be modeled using an ARIMA process. Further still, in each of the models above, the main κ_t time series was modelling with a random walk with drift. See for example Lee Carter (1992), Cairns, *et al.* (2006), Plat (2009). In other words, there is an implicit assumption in the forecasting element of these models that the drift is not going to change direction in the future. Given that long term forecasts are to be made with the model it is important that the direction of the drift is therefore reflective of the future direction of mortality improvement. It is this that we test in a wide range of scenarios in this paper.

For each of these models the approach in the literature has been to identify the best ARIMA process to fit to the period effect using a Bayesian Information Criterion approach. For example, in the case of Plat (2009) the approach of using the best resulting BIC measure on a selection of ARIMA(p,d,q) models results in the selecting an ARIMA(0,1,0) model for κ_t^1 and

ARIMA(1,0,0) for the remaining period effect factors κ_t^2 and κ_t^3 . This is common across all the models and as it pointed out in Plat(2009) it is the κ_t^1 factor that drives the main direction of mortality dynamics. Hence this is the parameter that needs to be modeled correctly.

3. METHODOLOGY

To identify if the period trend is changing over time we employ the method of structural breaks to identify deviations from the fitted trend. We first fit the models to the data and extract the corresponding time series κ_t or κ_t^1 (in the case of Cairns, Blake and Dowd, 2006; Plat, 2009; and O'Hare and Li, 2012) that represents the main period improvement factor. This time series reflects the average mortality rate improvement factor in each of the models and is the main driver of the forecasts of mortality derived from each of the models. We use Box-Jenkins approach to identify the most suitable ARIMA process to fit to the extracted κ_t which in all cases turns out to be a simple random walk with drift. If the selected ARIMA processes are appropriate then we should expect residuals whose mean does not deviate significantly from zero. We apply the tests of Bai and Perron (1998, 2003) to identify and date the structural breaks present. For the purposes of this paper we look at all the period effects for the multi factorial models but as we noted earlier the main driver of mortality forecasts based on any of the above models is that of the main period effect κ_t^1 . In the following subsection we describe the Bai and Perron method.

3.1. Bai and Perron's method. The idea of Bai and Perron (1998, 2003) estimating break points is based on calculating the global minimisers of the sum of squared residuals from say, a regression model. Now consider the piecewise linear regression method to compute the global minimisers of the objective function. The piecewise linear function begins with an ordering of the observations and applies a dynamic program. With a sample of size T , the total number of possible segments is at most $T(T + 1)/2$. The global sum of squared residuals for any m -partition and for any value of m must necessarily be a particular linear combination of these $T(T + 1)/2$ sums of squared residuals. The estimates of the break dates, the m -partition $(\hat{T}_1, \dots, \hat{T}_m)$, correspond to this linear combination which must have a minimal value of sum of squared residuals, so that an optimal partition can be selected over all possible m -partitions. Applying OLS segment by segment without constraints among them, once the sums

of squared residuals of the relevant partitions have been computed, a dynamic programming algorithm can be used to compare possible combinations, and then find a partition which achieves a global minimization over the entire sum of squared residuals. This method essentially proceeds through a sequential examination of optimal one-break partitions. Based on a recursive procedure. It simply minimizes sum of squared residuals over the set of possibilities obtained by assuming the first j elements have been optimally partitioned in $m - 1$ parts and that $m + 1$ part consists of the last $T - j + 1$ elements. The optimal partitions can be obtained iteratively.

Bai and Perron's method is mostly used in the literature in identifying structural breaks, recently Zeileis *et al.* (2003) developed the *R* package *strucchange* which includes most of the existing procedures dealing with structural breaks and their latest developments, for instance Zeileis (2000, 2005).

3.2. Fitting quality. In order to test the impact of structural breaks we will carry out some fitting and backtesting without allowing for the structural breaks and then, in the cases where structural breaks are identified, in the presence of those structural breaks. In each case we will compare the models with and without structural breaks allowed for using a selection of error measures. We use the Mean average percentage error (E1), the mean absolute percentage error (E2) and the Root mean square error (E3). The definitions of these are set out below for information, where X_1 and X_2 are the lower and upper limits of the age range considered, and $1, \dots, T$ are the years of data considered.

The average error, $E1$ – this equals the average of the standardized errors,

$$(5) \quad E1 = \frac{1}{X_1 - X_2 + 1} \sum_{x=X_1}^{X_2} \sum_{t=1}^T \frac{\text{projected}(m_{x,t}) - \text{actual}(m_{x,t})}{\text{projected}(m_{x,t})},$$

this is a measure of the overall bias in the projections. The average absolute error, $E2$ – this equals the average of absolute value of the standardized errors,

$$(6) \quad E2 = \frac{1}{X_1 - X_2 + 1} \sum_{x=X_1}^{X_2} \sum_{t=1}^T \left| \frac{\text{projected}(m_{x,t}) - \text{actual}(m_{x,t})}{\text{projected}(m_{x,t})} \right|,$$

this is a measure of the magnitude of the differences between the actual and projected rates. The standard deviation of the error, $E3$ – this equals the square root of the average of the squared

errors,

$$(7) \quad E3 = \sqrt{\frac{1}{X_1 - X_2 + 1} \sum_{x=X_1}^{X_2} \sum_{t=1}^T \left(\frac{\text{projected}(m_{x,t}) - \text{actual}(m_{x,t})}{\text{projected}(m_{x,t})} \right)^2}.$$

4. DATA

The data that we use in this paper comes from the Human Mortality Database.³ The data available for each country includes number of deaths $D_{x,t}$ and exposure to death $E_{x,t}$ for lives aged x last birthday during year t , and we use this to gain a proxy for the central mortality rate for lives aged x during year t as:

$$(8) \quad m_{x,t} = \frac{D_{x,t}}{E_{x,t}}.$$

The data provides an estimate of the true mortality due to issues with the recording of data. Death data tends to be recorded accurately, with death certificates in most cases. However, exposure data is taken from census data which may only be accurately recorded every 5 or 10 years adjusting these figures for migration, deaths and births etc. The resulting mortality estimates are therefore quite noisy, particularly at the older ages where there is less data available. The estimated nature of mortality data may be of concern when seeking to find structural breaks as noted by Cannon (2010). However, in this analysis we are testing the extracted time series rather than the data itself and it is this time series that drives the dynamics of mortality rates. It is the direction of this time series that requires closer inspection.

Data is available going back to the mid nineteenth century in some cases but we have restricted this study to data from 1960-2009 in order to have a consistent period across all countries and we have considered the following 30 countries in table 1 in this study. The period from 1960 - 2009 has also been used in many mortality studies using models that we are comparing see for example, Cairns *et al.* (2006, 2009), Plat (2009), O'Hare and Li (2012).

The wide range of countries give a good spread of populations both geographically and in terms of economic development. The inclusion of Male and Female data also enables gender differences to be considered. We focus on the age range 20-89 for several reasons. Firstly, the models upon which we have based our comparisons are also fitted to this age range. Secondly,

³This can be found at <http://www.mortality.org/> and was accessed in June 2013. The database is maintained in the Department of Demography at the University of California, Berkeley, USA, and at the Max Planck Institute for Demographic Research in Rostock, Germany.

TABLE 1. Countries considered in this study along with HMD codes

HMD Code	Country	HMD Code	Country
nor	Norway	pol	Poland
fin	Finland	usa	United States of America
lit	Lithuania	por	Portugal
spa	Spain	ukr	Ukraine
ast	Austria	czr	Czechoslovakia
fra	France	ity	Italy
swe	Sweden	rus	Russia
blr	Belarus	den	Denmark
nth	Netherlands	jap	Japan
swi	Switzerland	svk	Slovakia
bel	Belgium	est	Estonia
hun	Hungary	lat	Latvia
nzd	New Zealand	ukt	United Kingdom
bul	Bulgaria	uks	Scotland
can	Canada	uke	England

and as identified by Currie (2011), data at the older ages provide additional problems in terms of the reliability. Indeed in several cases mortality rates determined using older data appear to fall sharply beyond age 95.

5. EMPIRICAL ANALYSIS

In the following section we formalise our tests for structural breaks in the time series that we have extracted from each of the models in the previous section. This section is broken down as follows; in the first part we test the period effects of each model, using data from each country (both male and female) and present the break dates that are found in the analysis at a 95% confidence level. We also present some of the residual test plots carried out in the process of this analysis.

Having identified the presence of break points we consider the question of fitting quality allowing for those break points. We fit the models without allowing for break points and then fit it again, allowing for break points, by excluding data prior to the break point in the fitting process. We assess the fitting quality using the three measures E1, E2 and E3 as outlined above. Finally, we forecast the models with and without break points allowed for in sample (using the data from 1960 - 2000 and forecasting up to 2009) and out of sample (forecasting to 2019).

5.1. Identifying structural breaks. To identify the structural breaks we use the methods of Bai and Perron (1998, 2003). Specifically, having extracted the residuals from the fitted main

period effects for each of the models for both males and females, we plot the residuals along with the breakpoint test results. The figure 1 and figure 2 shows the results in the case of the Lee-Carter model for males and females for each of the 30 countries.

As can be seen from the figures, break dates can be identified in almost all cases with some level of confidence. In tables 2 through to 5 we show numerically the identified break dates for the main period effect factor κ_t and for the additional period effect parameters in the multi factor cases, at a 95% confidence level. In the tables we can see that the presence of breakpoints is prevalent and so should not be ignored.

TABLE 2. Break date results for the period effect of the Lee-Carter model for Males and Females.

Country	Male κ_t	Female κ_t	Country	Male κ_t	Female κ_t
Norway	1987	NA	Poland	NA	NA
Finland	1969	NA	USA	1967	NA
Lithuania	NA	NA	Portugal	NA	NA
Spain	NA	NA	Ukraine	NA	NA
Austria	1969	1968	Czechoslovakia	1988	1984
France	1968	NA	Italy	1983	NA
Sweden	1977	NA	Russia	1986	NA
Belarus	NA	NA	Denmark	1993	NA
Netherlands	1972	NA	Japan	NA	NA
Switzerland	NA	NA	Slovakia	1990	NA
Belgium	NA	NA	Estonia	1985	1992
Hungary	1993	NA	Latvia	1985	NA
New Zealand	1986	NA	UK	NA	NA
Bulgaria	NA	NA	Scotland	1991	NA
Canada	1974	NA	England	1983	NA

We note that the presence of structural breaks in the fitting of the period effects is significant but not conclusive. Given that there appears to be no pattern to whether or not a structural break will be present, the modelling process when fitting time series to period effects for a long term needs careful consideration. There are some comments that can be made about the presence of structural breaks at this stage:

- Firstly, it is clear that in the simpler Lee-Carter model the period effect shows a structural break in the case of Male data more often than in the case of female data. The period effect captures the improvement in mortality rates that can in some way be associated to all ages. The fact that breaks appear present in male data more than in female suggests that improvement rates has changed for males more than for females at least

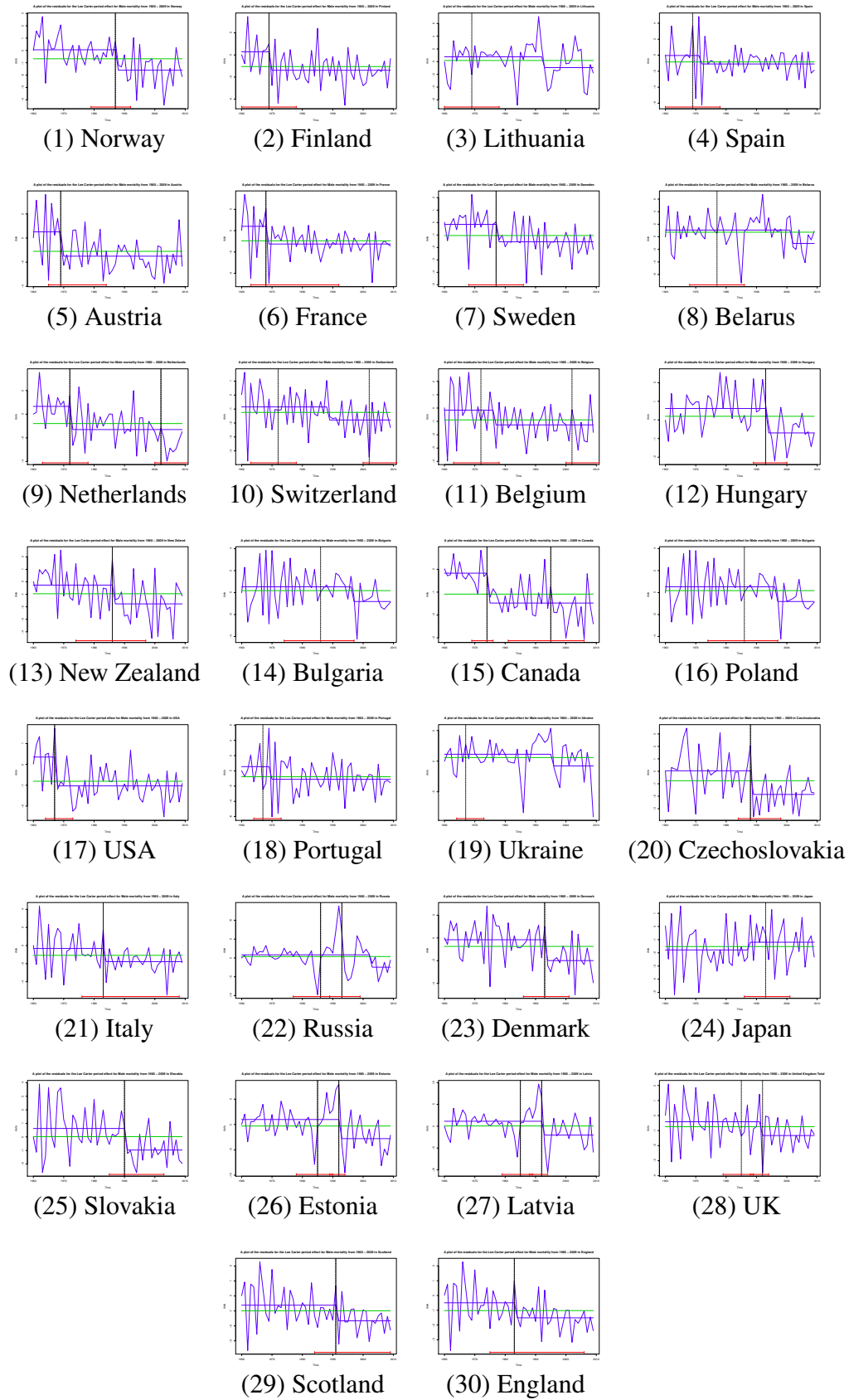


FIGURE 1. Structural Break plots for country specific period effects κ_t using the Lee-Carter model fitted to male data from 1960 - 2009

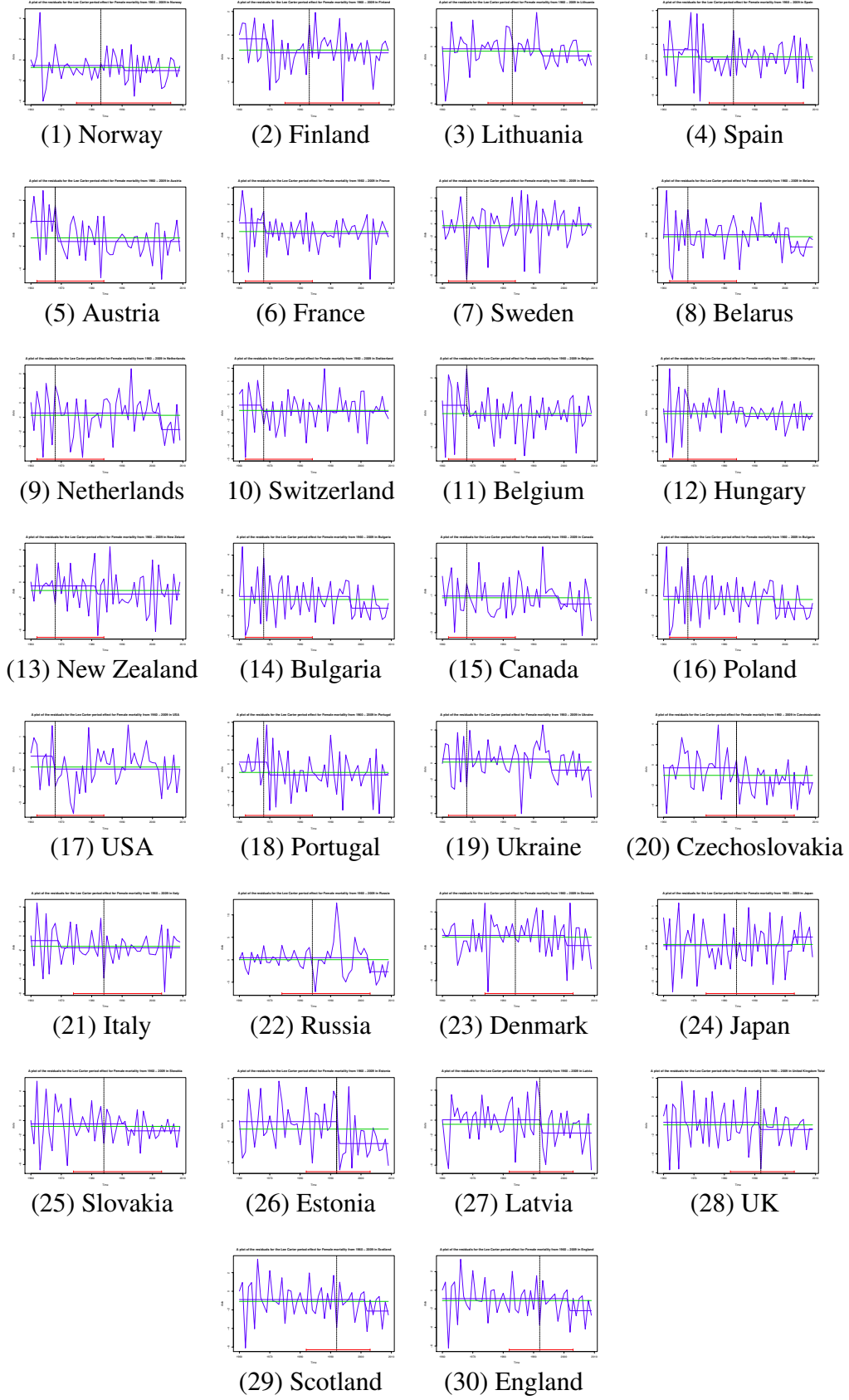


FIGURE 2. Structural Break plots for country specific period effects κ_t using the Lee-Carter model fitted to female data from 1960 - 2009

TABLE 3. Break date results for the period effect of the CBD model for Males and Females.

Country	Male		Female		Country	Male		Female	
	κ_t^1	κ_t^2	κ_t^1	κ_t^2		κ_t^1	κ_t^2	κ_t^1	κ_t^2
Norway	1987	NA	NA	NA	Poland	1991	NA	NA	NA
Finland	1971	NA	NA	NA	USA	1967	NA	1971	NA
Lithuania	NA	NA	NA	NA	Portugal	NA	NA	NA	NA
Spain	1996	1995	NA	NA	Ukraine	NA	NA	NA	NA
Austria	1982	NA	1968	NA	Czechoslovakia	1988	NA	1988	NA
France	1968	NA	NA	NA	Italy	1991	NA	NA	NA
Sweden	1977	NA	NA	NA	Russia	1986	1979	NA	NA
Belarus	NA	NA	NA	NA	Denmark	1993	NA	1993	NA
Netherlands	1972	NA	2002	NA	Japan	NA	NA	1987	1980
Switzerland	1989	NA	NA	NA	Slovakia	1990	NA	NA	NA
Belgium	1978	NA	NA	NA	Estonia	1985	NA	1992	NA
Hungary	1993	NA	1993	NA	Latvia	1985	1985	NA	NA
New Zealand	1978	NA	NA	NA	UK	NA	NA	NA	NA
Bulgaria	NA	NA	NA	NA	Scotland	1977	NA	NA	NA
Canada	1974	NA	NA	NA	England	1977	NA	NA	NA

during the period considered in this paper. It is important to note when considering the period effect that a visual observation of a plot of κ_t shows us that any change in improvement has been to accelerate life expectancy improvements beyond those of previous years. In other words, if this observation is not captured in the model then we are at risk of underestimating life expectancy.

- Looking to the CBD (2006) model, we test both period effects in the simple model (without cohort effect additions) and conclude with similar comments that males are more impacted by the structural break issue than females. We also note that the presence of breaks is more prevalent in the κ_t^1 factor than in the κ_t^2 factor.
- The introduction of multiple linear factors as in the Plat (2009) model increases the observations of structural breaks in each of the period effects and also increase the presence of breaks in the female model.
- The O'Hare and Li (2012) model still results in many observations of structural breaks but many fewer than its nearest model, the Plat (2009) model and fewer in the κ_t^3 factor which is the innovation in that model. It being linked to a quadratic age effect. If we regard structural breaks as a simple case of non-linearity, our results indicate that non-linearity should be taken into account in mortality modelling. Testing of non-linearity

TABLE 4. Break date results for the period effect of the Plat model for Males and Females.

Country	Male			Female		
	κ_t^1	κ_t^2	κ_t^3	κ_t^1	κ_t^2	κ_t^3
Norway	1987	1987	1989	1988	1987	NA
Finland	1971	NA	NA	NA	1987	1988
Lithuania	1992	1993	NA	1993	1980	NA
Spain	NA	1982	1982	NA	1987	1976
Austria	1982	1983	NA	1968	1983	NA
France	1968	1976	NA	1968	1976	NA
Sweden	1977	1980	1983	1990	1980	NA
Belarus	NA	NA	NA	2001	NA	1985
Netherlands	1970	NA	NA	2002	1979	NA
Switzerland	1993	NA	NA	NA	1977	NA
Belgium	NA	NA	NA	NA	1977	NA
Hungary	1993	1994	NA	1993	1993	NA
New Zealand	1974	NA	NA	NA	NA	NA
Bulgaria	1993	1993	1985	1996	1991	1990
Canada	1977	NA	NA	NA	1979	NA
Poland	1991	1991	1987	1991	1988	1988
USA	1967	1991	NA	1969	1969	1982
Portugal	NA	NA	NA	1993	1986	1986
Ukraine	NA	NA	1987	1995	NA	1986
Czechoslovakia	1988	1988	1986	1988	1990	1985
Italy	1983	NA	NA	1969	1969	1982
Russia	NA	NA	NA	2002	1993	1986
Denmark	1987	1987	NA	1993	1968	NA
Japan	NA	NA	1987	NA	1972	1970
Slovakia	1990	NA	NA	1991	NA	NA
Estonia	1992	NA	NA	1992	1992	NA
Latvia	1992	NA	NA	1992	1992	NA
UK	1977	NA	NA	1976	NA	NA
Scotland	1977	NA	1985	1977	1967	1985
England	1984	NA	1985	1983	1967	1982

in its simplest form, structural break, is a first step to have a better nonlinear model of mortality in the future.

Given that in each of these cases we have not changed the data sets, what we are identifying here is a mis-specification of the model structures rather than a comment on the data. However, some insights on the data can be inferred, particularly why the impact appears overall to be greater for males than for females.

5.2. Fitting the models with and without allowance for structural breaks. The next stage in the analysis is to fit the models allowing for the observed structural breaks. We consider each of the 30 countries covered in the paper again, and this time fit the model to data from 1960 - 2009

TABLE 5. Break date results for the period effect of the OL model for Males and Females.

Country	Male			Female		
	κ_t^1	κ_t^2	κ_t^3	κ_t^1	κ_t^2	κ_t^3
Norway	1987	1987	1989	1988	1987	NA
Finland	1971	NA	NA	NA	NA	NA
Lithuania	NA	NA	NA	NA	1980	NA
Spain	NA	1995	1982	NA	1987	1976
Austria	1982	1983	NA	1968	1983	NA
France	1968	1976	NA	1968	NA	NA
Sweden	1977	1975	1980	NA	1972	NA
Belarus	NA	NA	NA	2001	NA	NA
Netherlands	1970	NA	1987	2002	1977	NA
Switzerland	1993	NA	NA	NA	1977	NA
Belgium	NA	NA	NA	NA	1977	NA
Hungary	1993	1994	NA	1993	NA	NA
New Zealand	1974	NA	NA	NA	NA	NA
Bulgaria	1996	NA	NA	1996	NA	1990
Canada	1977	1977	NA	NA	NA	NA
Poland	1991	NA	NA	1991	NA	NA
USA	1967	1991	NA	1967	NA	NA
Portugal	NA	NA	NA	NA	1982	1986
Ukraine	NA	NA	NA	NA	NA	1983
Czechoslovakia	1988	1988	1982	1988	1990	NA
Italy	1983	NA	NA	1969	NA	NA
Russia	NA	NA	NA	NA	NA	1993
Denmark	1987	1987	NA	1993	1969	NA
Japan	NA	NA	1979	NA	NA	1996
Slovakia	1990	NA	NA	NA	NA	NA
Estonia	1992	NA	NA	1992	1992	NA
Latvia	1985	NA	NA	NA	NA	NA
UK	1977	NA	NA	NA	NA	NA
Scotland	1977	NA	1982	1977	NA	1982
England	1977	NA	1979	1977	NA	1982

with and without the allowance for structural breaks. We present the results in tables including all countries but where we only populate the cases where a structural break was present. We present results below in tables 6 - 8 using the three measures of error E1, E2 and E3 outlines earlier.

TABLE 6. Fitting results (expressed as percentages) measured using the mean absolute percentage error (E1) for Males and Females for the Lee-Carter, CBD, Plat, and OL models.

Country	Lee-Carter						CBD						Plat						OL					
	Male			Female			Male			Female			Male			Female			Male			Female		
	W/out	With	W/out	W/out	With	With	W/out	With	W/out	W/out	With	With	W/out	With	W/out	W/out	With	With	W/out	With	W/out	With	With	With
Norway	8.34	8.10					18.69	18.69					7.49	7.25	10.93	10.76			45.14	7.36	11.10	10.69		
Finland	7.70	7.63					12.53	12.53					6.39	6.37					55.32	6.35				
Lithuania																								
Spain							15.20	15.17																
Austria	6.27	6.06	8.19	8.12			14.14	14.15	23.90	23.88			5.95	5.79	7.89	7.95			49.73	5.76	8.11	8.14		
France	5.44	7.21					13.83	13.83					5.93	6.01	6.40	6.48			51.96	6.09	6.56	6.59		
Sweden	6.60	6.47					16.62	16.62					6.48	6.44	9.18	8.45			43.40	6.24				
Belarus																								
Netherlands	5.63	5.27					13.99	13.99	20.29	19.04			4.56	4.53	6.65	6.11			41.02	4.68	7.14	6.24		
Switzerland							20.17	20.21					7.36	7.02					46.07	6.99				
Belgium							15.30	15.29																
Hungary	17.12	4.93					20.92	20.97	14.89	14.82			7.01	5.93	7.77	6.62			52.93	6.68	7.82	7.24		
New Zealand	9.49	8.37					21.67	21.67					9.43	9.35					40.14	9.51				
Bulgaria													4.73	4.65	7.64	6.66			48.03	4.53	7.39	6.53		
Canada	5.08	4.96					15.57	15.59					4.08	3.95					45.46	4.04				
Poland							7.07	7.09					2.86	2.47	4.17	3.84			57.25	2.63	5.34	4.37		
USA	4.43	4.37					13.57	13.58	11.71	11.62			2.61	2.70	2.95	2.92			45.63	2.71	2.98	2.99		
Portugal															13.85	13.37								
Ukraine															5.60	4.58								
Czechoslovakia	5.64	4.97	7.36	7.17			11.00	11.01	16.72	16.66			4.97	4.84	7.06	6.69			51.47	4.68	7.41	6.68		
Italy	6.64	6.07					17.85	17.89					4.21	4.03	4.14	4.09			50.03	4.20	4.38	4.25		
Russia	5.14	4.76					6.66	6.66							4.91	2.25								
Denmark	8.35	7.61					12.77	12.77	13.16	13.17			7.58	7.47	11.40	11.38			41.16	7.55	11.57	11.35		
Japan									27.38	27.15														
Slovakia	11.19	5.94					10.42	10.44					5.74	5.50	10.21	9.88			55.27	5.45				
Estonia	11.61	9.91	21.36	19.42			11.20	11.21	22.02	21.94			10.00	9.69	19.21	18.85			60.71	9.63	19.19	18.85		
Latvia	10.27	6.83					8.18	8.18					5.59	5.37	9.00	8.21			58.29	5.78				
UK													4.21	3.98	3.48	3.36			41.28	4.21				
Scotland	8.17	6.78					15.67	15.66					7.91	7.72	9.13	9.13			42.58	8.13	9.35	9.27		
England	5.38	4.00					15.41	15.39					4.60	3.89	3.82	3.47			40.97	4.23	4.00	3.69		

TABLE 7. Fitting results (expressed as percentages) measured using the mean average percentage error (E2) for Males and Females for the Lee-Carter, CBD, Plat, and OL models.

Country	Lee-Carter						CBD						Plat						OL					
	Male			Female			Male			Female			Male			Female			Male			Female		
	W/out	With	W/out	W/out	With	With	W/out	With	W/out	W/out	With	With	W/out	With	W/out	W/out	With	With	W/out	With	W/out	W/out	With	With
Norway	2.03	1.56					-7.26	-7.24					1.25	1.24	2.02	2.58			-45.14	1.33	2.57	2.62		
Finland	1.39	1.17					-3.98	-3.98					0.75	0.80					-55.32	0.87				
Lithuania													0.44	0.46	1.90	1.99								
Spain							-7.42	-7.34																
Austria	0.50	0.76	1.73	1.54			-5.61	-5.59	-14.90	-14.87			0.45	0.42	0.92	1.12			-49.73	0.51	1.18	1.22		
France	0.81	4.85					-1.25	-1.24					4.13	4.23	4.60	4.77			-51.96	4.20	4.53	4.58		
Sweden	1.01	0.87					-9.38	-9.36					0.79	0.86	0.77	1.64			-43.40	0.94				
Belarus																								
Netherlands	1.25	0.38					-3.07	-3.07	-10.96	-10.14			0.54	0.53	0.85	0.75			-41.02	0.54	0.07	0.66		
Switzerland							-8.94	-8.90					1.30	1.23					-46.07	1.07				
Belgium							-5.96	-5.96																
Hungary	13.98	0.61					13.02	13.04	-2.25	-2.08			1.91	0.02	2.55	0.71			-52.93	0.08	0.40	0.54		
New Zealand	1.55	1.20					-2.13	-2.13					1.85	1.74					-40.09	1.73				
Bulgaria													0.46	0.34	1.78	1.27			-48.03	0.43	1.06	1.17		
Canada	0.72	0.48					-3.62	-3.62					0.49	0.55					-45.46	0.53				
Poland							0.41	0.41					0.31	0.11	0.36	0.43			-57.25	0.12	0.18	0.52		
USA	0.55	0.63					-3.02	-3.02	-6.40	-6.36			0.25	0.36	0.20	0.27			-45.63	0.36	0.29	0.25		
Portugal															11.00	10.87								
Ukraine															1.16	0.76								
Czechoslovakia	0.55	0.61	1.47	1.27			-0.59	-0.59	-10.66	-10.58			0.34	0.46	0.74	1.13			-51.47	0.52	1.05	1.13		
Italy	1.22	1.20					-8.36	-8.32					0.53	0.56	0.26	0.43			-50.03	0.59	0.35	0.38		
Russia	0.98	0.43					-1.62	-1.63							1.05	0.26								
Denmark	1.94	1.51					-5.90	-5.87	-0.69	-0.46			1.36	1.12	2.37	3.03			-41.16	1.15	2.57	3.00		
Japan									-16.81	-16.67														
Slovakia	7.34	0.66					2.93	2.94					0.67	0.42	2.63	2.54			-55.27	0.43				
Estonia	4.23	1.54	5.60	8.22			0.65	0.64	-6.51	-6.36			1.82	1.77	8.23	8.13			-60.71	2.00	8.03	8.18		
Latvia	5.10	0.98					0.07	0.07					0.42	0.26	2.57	1.80			-58.29	0.60				
UK													0.62	0.67	0.11	0.36			-41.28	0.71				
Scotland	0.75	1.02					-2.29	-2.31					1.51	1.50	1.84	2.15			-42.58	1.62	2.31	2.27		
England	0.87	0.39					-4.60	-4.63					0.71	0.69	0.13	0.47			-40.97	0.71	0.50	0.47		

TABLE 8. Fitting results (expressed as percentages) measured using the root mean square percentage error (E3) for Males and Females for the Lee-Carter, CBD, Plat, and OL models.

Country	Lee-Carter						CBD						Plat						OL					
	Male			Female			Male			Female			Male			Female			Male			Female		
	W/out	With		W/out	With		W/out	With		W/out	With		W/out	With		W/out	With		W/out	With		W/out	With	
Norway	0.83	0.83					1.10	1.10					0.71	0.82		0.56	0.73		2.21	0.91		0.59	0.72	
Finland	0.95	0.95					1.12	1.12					0.86	1.01		0.78	0.86		1.98	0.99				
Lithuania													1.15	1.14										
Spain							1.41	1.46																
Austria	0.77	0.77	0.60	0.60			1.27	1.28	1.02	1.02			0.77	0.97	0.52	1.14			2.00	0.88	0.55	1.00		
France	0.57	0.73					1.14	1.15					0.77	1.14	0.66	1.09			1.88	1.08	0.68	1.00		
Sweden	0.62	0.62					0.97	0.98					0.59	0.78	0.49	0.50			1.98	0.76				
Belarus															1.25	1.34					1.25	1.38		
Netherlands	0.75	0.74					0.81	0.81	1.52	1.64			0.54	0.64	0.64	0.94			1.99	0.67	0.72	0.85		
Switzerland							1.34	1.37					0.67	0.96					2.18	0.88				
Belgium							1.08	1.08																
Hungary	1.90	1.20					1.34	1.35	1.35	1.38			0.90	0.83	0.62	0.83			2.41	0.91	0.64	0.72		
New Zealand	0.98	0.97					1.17	1.17					0.95	0.96					2.04	1.00				
Bulgaria													0.91	0.88	0.77	1.02			2.27	0.94	0.81	1.01		
Canada	0.72	0.71					0.87	0.88					0.54	0.70					1.99	0.75				
Poland							1.14	1.15					0.78	0.70	0.50	1.03			2.35	0.66	0.57	0.91		
USA	0.58	0.57					1.01	1.01	0.89	0.90			0.42	0.83	0.40	0.50			1.85	0.83	0.44	0.48		
Portugal															1.11	1.24								
Ukraine															0.68	0.65								
Czechoslovakia	0.96	0.94	0.71	0.71			1.14	1.15	1.10	1.11			0.88	0.88	0.61	0.91			2.24	0.86	0.65	0.87		
Italy	0.69	0.69					1.14	1.16					0.61	0.73	0.41	0.94			2.05	0.82	0.48	0.87		
Russia	1.12	1.12					1.55	1.55							0.83	0.84								
Denmark	0.85	0.83					1.09	1.09	0.91	0.94			0.79	0.99	0.65	0.74			2.20	0.93	0.67	0.72		
Japan									1.12	1.16														
Slovakia	1.37	1.13					1.20	1.21					0.98	0.98	0.76	0.81			2.33	1.00				
Estonia	1.47	1.38	1.04	0.97			1.44	1.44	1.43	1.46			1.20	1.18	0.96	1.02			2.59	1.21	0.99	1.00		
Latvia	1.24	1.20					1.41	1.41					1.09	1.15	0.79	0.75			2.38	1.07				
UK													0.55	0.94	0.40	0.60			2.08	1.02				
Scotland	0.89	0.87					1.01	1.01					0.80	0.84	0.59	0.60			2.14	0.94	0.61	0.62		
England	0.63	0.61					0.90	0.92					0.60	1.03	0.41	0.66			2.07	1.03	0.45	0.63		

We note from the above results that in terms of fitting quality, allowing for the structural break results in an improved fit almost exclusively. In most cases that improvement is minor but in some it can be more than 0.5%. Noting that the fitting quality is improved primarily because the direction of the trend has been updated. It also follows that the forecasts produced by the model, with and without this allowance for the structural break will differ, we leave the quantification of this for further research but note that time series extrapolative models in a demographic context should be used with caution.

5.3. In sample forecasting with and without structural breaks. Having identified the presence of structural breaks, it is important to consider the impact of these on forecasting. For illustration purpose, we will firstly consider in sample forecasting in the case of the Lee-Carter model. For this we fit the model to data across the 30 countries and for both male and female from 1960 - 2000 and then use the resulting parameter estimates to forecast from 2001-2009. The results of this in sample forecast for the Lee-Carter model are presented in table 9. It can be seen from table 9 that almost exclusively the allowance for a structural break improves the accuracy of the mortality forecast, in some cases significantly. This is mainly due to the fact that a more recent trend will better reflect future mortality rates and it is noted that the results are very varied with very minor improvement in some cases (see for example Italy E2 measure), but large improvements in other cases (see for example Hungary E2 measure). The size of the improvement if present is largely driven by the date of the structural change.

TABLE 9. Backtesting results (expressed as percentages) measured using the mean average percentage error (E1), the mean absolute percentage error (E2) and root mean square percentage error (E3) for Males and Females for the Lee-Carter model.

Country	E1 measure				E2 measure				E3 measure			
	Male		Female		Male		Female		Male		Female	
	W/out	With	W/out	With	W/out	With	W/out	With	W/out	With	W/out	With
Norway	28.21	22.88			-27.94	-22.62			1.98	1.91		
Finland	41.83	40.2			-41.6	-39.93			3.56	3.54		
Lithuania												
Spain												
Austria	43.6	39.96	71.59	67.33	-43.6	-39.96	-71.59	-67.33	3.48	3.33	4.24	4.01
France	34.46	32.08			-34.36	-31.77			3.37	3.27		
Sweden	40.8	37.8			-40.8	-37.8			2.76	2.7		
Belarus												
Netherlands	26.91	24.63			-26.89	-24.61			1.67	1.67		
Switzerland												
Belgium												
Hungary	43.03	66.16			24.43	60.43			2.73	2.32		
New Zealand	39.01	36.5			-35.53	-30.56			3.24	3.04		
Bulgaria												
Canada	41.67	40.05			-41.67	-40.05			2.44	2.47		
Poland												
USA	35.23	33.16			-35.23	-33.16			2.76	2.72		
Portugal												
Ukraine												
Czechoslovakia	33.8	23.93	65.2	64.11	-33.7	-23.46	-65.2	-64.11	3.31	2.86	4	3.92
Italy	37.49	36.15			-36.98	-36.15			3.22	3.03		
Russia	158.67	104.36			158.67	104.36			3.78	3.43		
Denmark	29.99	17.68			-29.57	-14.56			2.67	2.31		
Japan												
Slovakia	24.63	33.23			5.84	22.96			2.25	1.99		
Estonia	87.58	60.74	57.34	115.72	83.85	56.47	-57.29	10.62	2.69	2.55	3.59	5.19
Latvia	96.79	67.62			96.29	67			2.73	2.4		
UK												
Scotland	37.54	31.38			-22.66	-26.51			3.47	3.13		
England	36.95	36.81			-35.59	-36.6			3.32	3.19		

We now forecast our models allowing for and not allowing for structural breaks to see if the impact on future forecasts is significant. Figures 3-6 plot forecasts from 2010 to 2019 without adjustment of bias for the countries France, UK, Netherlands and USA for illustrative purposes. The results, and commentary are reflective of the full range of countries. Ignoring the basis issue the blue line reflects the forecast without allowing for the structural break whereas the red line allows for the change in the trend identified using the structural break approach. In all cases allowing for the trend change produces lower mortality rates and hence larger life expectancies although the results are varied with some cases showing clear differences in the forecast whilst other some less obvious differences. The same analysis has been completed for the multi factorial models and is available on request.

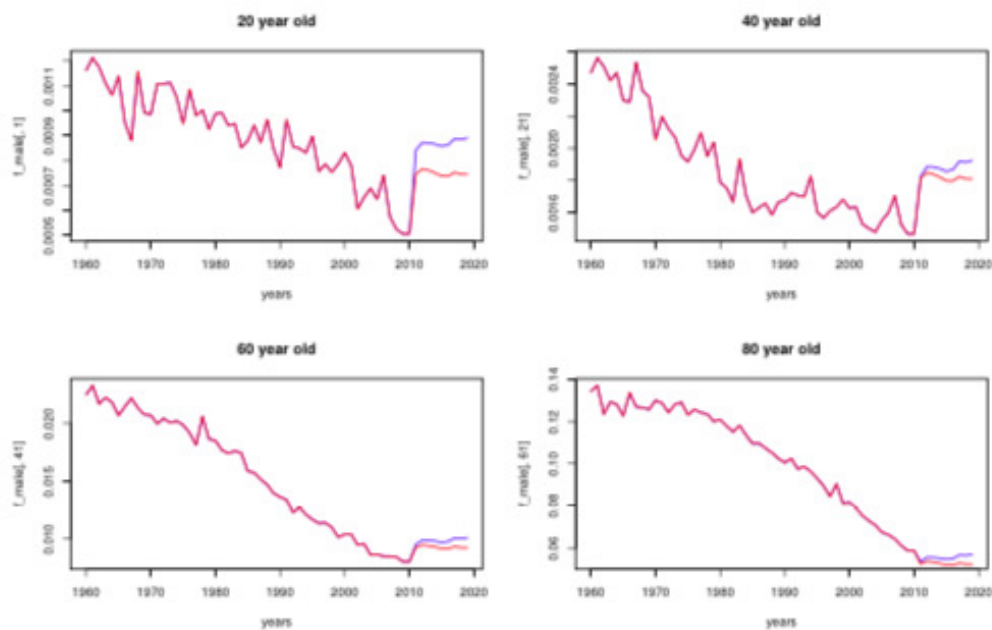


FIGURE 3. Forecasts to 2019 using the Lee-Carter model for France

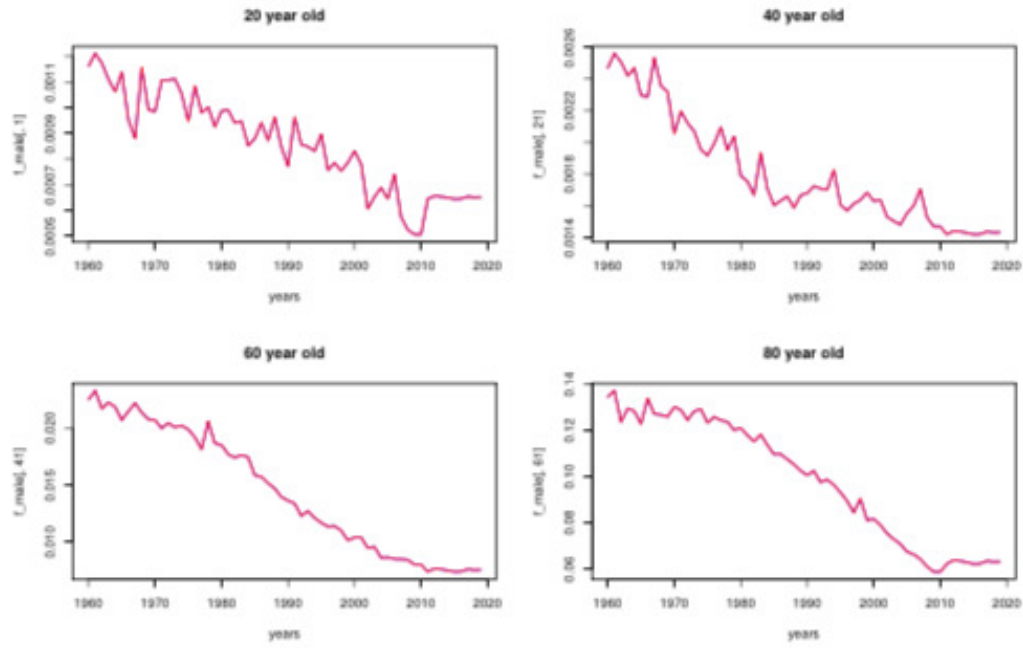


FIGURE 4. Forecasts to 2019 using the Lee-Carter model for the UK

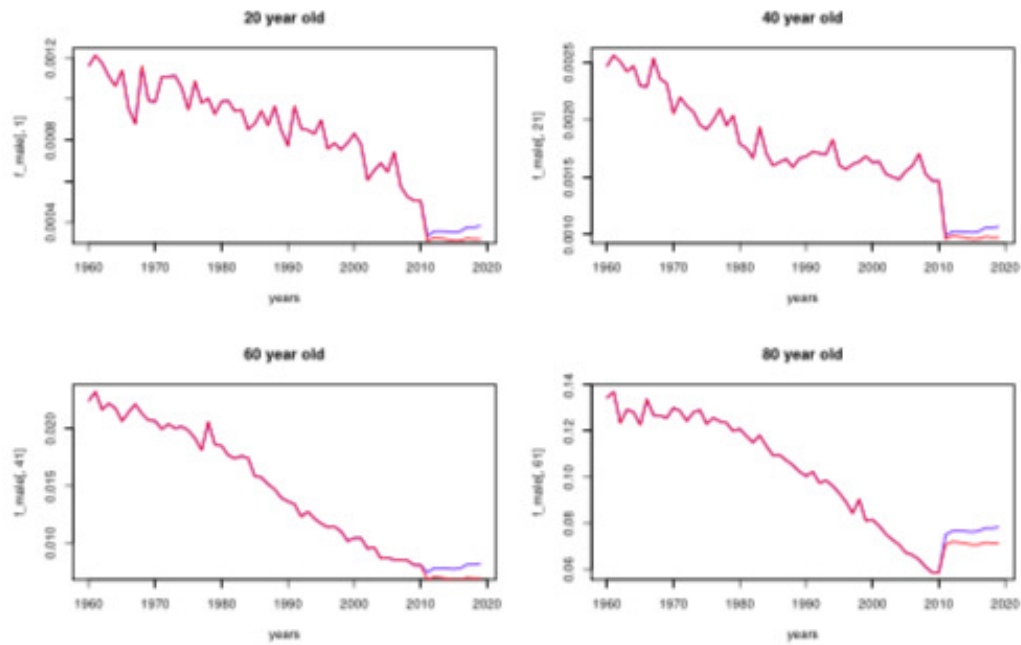


FIGURE 5. Forecasts to 2019 using the Lee-Carter model for the Netherlands

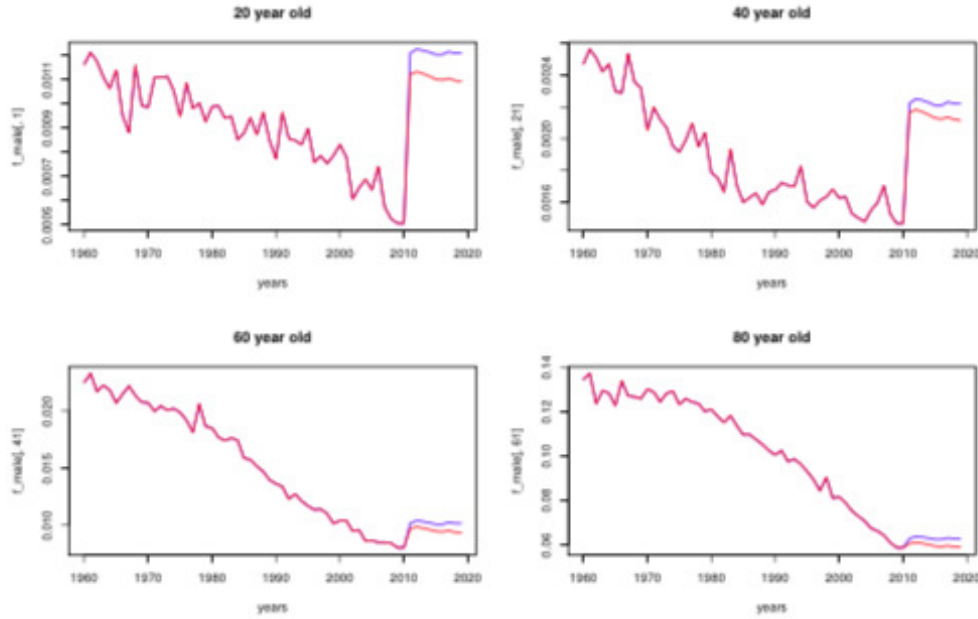


FIGURE 6. Forecasts to 2019 using the Lee-Carter model for the US

6. CONCLUSIONS

In this paper we have considered several of the leading extrapolative models of mortality rates and have applied the methods of Bai and Perron (2003) to test for the presence of structural breaks in the model specifications. More specifically we have fitted the models of Lee and Carter (1992), Cairns, Blake and Dowd (2006), Plat (2009) and O'Hare and Li (2012) to the data for 30 countries. Having noted that the forecasts of mortality resulting from these models are driven primarily by the κ_t (or κ_t^1) parameter, we have fitted the best ARIMA process to the extracted time series and then tested the residuals for deviation from zero.

In each case we found that there was indeed a breakpoint visible in the residuals falling somewhere around the 1970's confirming previous demographic research. We then carried out the forecasting process again making allowance for the structural break. The results show that in nearly two-thirds of cases the model allowing for structural breaks provides a more accurate fit measured on each of the $E1$, $E2$, and $E3$ measures. Whilst the findings are important in highlighting the importance of the sample period when fitting a model to mortality data they make no reference to future structural breaks. Further research could look at more recent developments in the identification of structural breaks in models. Namely, monitoring data for

structural breaks as and when they occur. This would then allow for these breaks to be incorporated into mortality models more efficiently reducing any future forecast errors. In addition, the allowance for structural breaks appears to make more of a difference for some models than others. Future research questions could investigate the reasons for this.

REFERENCES

- [1] Bai, J., and Perron, P., (1998). Estimating and Testing Linear Models With Multiple Structural Changes, *Econometrica*, 66, 47-78.
- [2] Bai, J., and Perron, P., (2003). Computation and Analysis of Multiple Structural Change Models, *Journal of Applied Econometrics*, 18, 1-22.
- [3] van Berkum, F., Katrien, A., and Vellekoop, M., (2014). The impact of multiple structural changes on mortality predictions, *FEBS Research Report AFI-1487*.
- [4] Booth, H., Maindonald, J., and Smith, L., (2002). Applying Lee-Carter under conditions of variable mortality decline, *Population Studies* 56, 325-336.
- [5] Booth, H., and Tickle, L., (2008). Mortality modeling and forecasting: A review of methods, *Annals of actuarial science*. 3, 3-43.
- [6] Brouhns, N., Denuit, M., and Vermunt, J.K., (2002). A Poisson log-bilinear approach to the construction of projected lifetables, *Insurance: Mathematics and Economics* 31(3), 373-393.
- [7] Brown R.L., Durbin J., and Evans J.M., (1975). Techniques for testing constancy of regression relationships over time, *Journal of the Royal Statistical Society, Series B*, 37, 149-163.
- [8] Cairns, A.J.G., Blake, D., and Dowd, K., (2006). A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration, *Journal of Risk and Insurance* 73, 687-718.
- [9] Cairns, A.J.G., Blake, D., Dowd, K., Coughlan, G.D., Epstein, D., Ong, A., and Balevich, I., (2009). A quantitative comparison of stochastic mortality models using data from England & Wales and the United States, *North American Actuarial Journal* 13(1), 1-35.
- [10] Cairns, A.J.G., Blake, D., Dowd, K., Coughlan, G.D., Epstein, D., and Khalaf-Allah, M., (2011). Mortality density forecasts: An analysis of six stochastic mortality models. *Insurance: Mathematics and Economics* 48, 355-367.
- [11] Cannon, E., (2010). Estimation and pricing with the Cairns-Blake-Dowd model of mortality, Pensions Institute Discussion Paper PI-1003.
- [12] Carter, L. R., and Prskawetz, A., (2001). Examining Structural Shifts in Mortality Using the Lee-Carter Method, *Working Paper WP 2001-007*, Max Planck Institute for Demographic Research
- [13] Clayton, D. & Schifflers, E., (1987). Models for temporal variation in cancer rates. II: Age-period-cohort models, *Statistics in Medicine*, 6, 469-481.
- [14] Coelho, E. and Nunes, L. C., (2011). Forecasting mortality in the event of a structural change, *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 174, 713-736.
- [15] Currie, I.D., (2006). Smoothing and forecasting mortality rates with P-splines, *Presentation to the Institute of Actuaries*. Available at: <http://www.ma.hw.ac.uk/iain/research.talks.html>.
- [16] Currie, I.D., (2011). Modelling and forecasting the mortality of the very old, *ASTIN Bulletin*, 41, 419-427.

- [17] Delwarde, A., Denuit, M., and Eilers, P., (2007). Smoothing the Lee & Carter and Poisson log-bilinear models for mortality forecasting: A penalized log-likelihood approach, *Statistical Modelling*, 7, 29-48.
- [18] Girosi, F., and King, G., (2005). *A reassessment of the Lee-Carter mortality forecasting method*, Working Paper, Harvard University.
- [19] Haberman, S., and Renshaw, A., (2011). A comparative study of parametric mortality projection models, *Insurance: Mathematics and Economics*, 48, 35-55.
- [20] Haberman, S., and Renshaw, A., (2012). Parametric mortality improvement rate modelling and projecting, *Insurance: Mathematics and Economics*, 50, 309-333.
- [21] Haberman, S., and Renshaw, A., (2013). Modelling and projecting mortality improvement rates using a cohort perspective, *Insurance: Mathematics and Economics*, 53, 150-168.
- [22] Harris, D., Harvey, D. I., Leybourne, S. J. and Taylor, A. M. R., (2009). Testing for a unit-root in the presence of a possible break in trend, *Econometric Theory*, 25, 1545-1588.
- [23] Harvey, D. I., Leybourne, S. J. and Taylor, A. M. R., (2009). Simple, robust and powerful tests of the changing trend hypothesis, *Econometric Theory*, 25, 995-1029.
- [24] Koissi, M.C., Shapiro, A.F., and Hognas, G., (2005). Evaluating and Extending the Lee-Carter Model for Mortality Forecasting: Bootstrap Confidence Interval, *Insurance: Mathematics and Economics*, 38, 1-20.
- [25] Lee, R.D. & Carter, L.R., (1992). Modeling and Forecasting U. S. Mortality, *Journal of the American Statistical Association*, 87(419), 659-671.
- [26] Lee, R. D., and Miller, T., (2001). Evaluating the performance of the Lee-Carter method for forecasting mortality, *Demography*, 38, 537-549.
- [27] Li, J.S.H., Chan, W.S. and Cheung, S.H., (2011). Structural Changes in the Lee-Carter Mortality Indexes: Detection and Implications, *North American Actuarial Journal*, 15, 13-31.
- [28] Mitchell, D., Brockett, P., Mendoza-Arriaga, R., and Muthuraman, K., (2013). Modeling and forecasting mortality rates, *Insurance: Mathematics and Economics*, 52, 275-285.
- [29] O'Hare, C., and Li, Y., (2012). Explaining young mortality, *Insurance: Mathematics and Economics*, 50(1), 12-25.
- [30] O'Hare, C., and Li, Y., (2014). Identifying structural breaks in mortality modelling, *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems: Part B. Mechanical Engineering*, forthcoming.
- [31] Plat, R., (2009). On stochastic mortality Modeling, *Insurance: Mathematics and Economics*, 45(3), 393-404.
- [32] Renshaw, A.E., Haberman, S., (2003). Lee-Carter mortality forecasting with age-specific enhancement, *Insurance: Mathematics and Economics*, 33, 255-272.
- [33] Renshaw, A.E., Haberman, S., (2006). A cohort-based extension to the Lee-Carter model for mortality reduction factors, *Insurance: Mathematics and Economics*, 38, 556-70.
- [34] Tuljapurkar, S., and Boe, C. (1998). Mortality Change and Forecasting: How Much and How Little Do We Know? *North American Actuarial Journal*, 2, 13-47.

- [35] Zeileis, A., (2000). p -values and alternative boundaries for CUSUM tests, *Working Paper 78, SFB Adaptive Information Systems and Modelling in Economics and Management Science*, <http://www.wu-wien.ac.at/am/wp00.htm#78>.
- [36] Zeileis, A., Kleiber C., Kramer W., Hornik K., (2003). Testing and Dating of Structural Changes in Practice, *Computational Statistics and Data Analysis*, 44, 109-123.
- [37] Zeileis, A., (2005). A Unified Approach to Structural Change Tests Based on ML Scores, F Statistics and OLS Residuals, *Econometric Reviews*, 24, 445-466.
- [38] Zivot, E., and Andrews, D., (1992). Further Evidence of the Great Crash, the Oil-Price Shock and the Unit-Root Hypothesis, *Journal of Business and Economic Statistics*, 10, 251-270.